

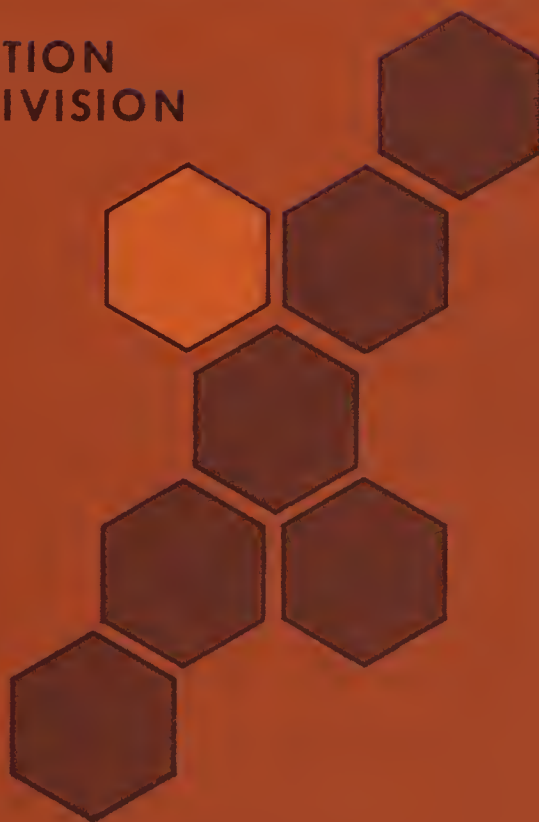
Historic, archived document

Do not assume content reflects current scientific knowledge, policies, or practices.

FPED WORKING PAPER

Reserve
aHD1447
.H37
1971

FARM PRODUCTION
ECONOMICS DIVISION



ECONOMIC RESEARCH SERVICE

U.S. DEPARTMENT OF AGRICULTURE

The manuscript has been reproduced for information and discussion with Farm Production Economics Division. The manuscript has not been cleared for publication and should not be cited as a reference. The views expressed are those of the author and do not necessarily represent the opinion of FPED, the Economic Research Service or the U.S. Department of Agriculture.



WP 7-72

U.S.D.A., NAL

AUG 14 2001

Cataloging Prep

A PRICE AND YIELD ENVIRONMENTAL SIMULATOR
BASED ON HISTORICAL DATA

by Virden L. Harrison and W.H.M. Morris*

Applied in 1970

2

A PRICE AND YIELD ENVIRONMENTAL SIMULATOR
BASED ON HISTORICAL DATA

by Virden L. Harrison and W. H. M. Morris *

In studies of farm and non-farm firms involving more than one year in the analysis, the problem of obtaining an appropriate series of prices and yields for all enterprises is acute. This paper describes (1) a method of establishing a price, yield, and cost environment for projecting into the future the historical relation among as many crop and livestock enterprises as are desired and (2) a formula by which a second set of price, yield, and cost data may be obtained which deviates from the first set with a specified statistical precision. 1

THE ENVIRONMENTAL SIMULATOR

The general concept is to develop a means of obtaining data, for as many years in the future as desired, such that (1) all cycles, market fluctuations, technology trends, and other time phenomena which appeared in the prices and yields in recent history, will still appear in the data, (2) the prices, yields, and costs among all crop and livestock enterprises chosen will remain in the actual historical relation one to another year after year, and (3) the data are in a form in which individual values may be changed to reflect current or expected conditions and trends in any or all of the enterprises.

Data meeting the above criteria are useful in almost all dynamic studies of farm management, including studies of enterprise choice over time, studies of growth strategies and their effect as enterprise conditions change over time, or in any study in which the decision process requires a change in enterprise combinations as a result of

a change in the expected relative profitability of each.

The basic approach is as follows: Theoretically, if the time trend² (if any) is taken out of data which cover the length of one complete price cycle, then the value in the first year of the cycle should be identical to the value the year after the end of the cycle, regardless of the stage of the cycle at its beginning (i.e. if the beef cattle cycle is 11 years long and the time trend removed, the value in year 1 is the same as the value in year 12; the value in year 2 is the same as in year 13; etc.). Thus the data for an enterprise can be extended beyond the latest year. Since the price cycle for beef cattle (about 11 years) may include about 3 hog cycles (about 3-4 years each), the length of the beef cycle may be chosen as a base period for all enterprises. The procedure, then, is to consider all enterprises during the base period, isolate from the data the individual time trend appropriate to each enterprise, extend the data for as many years as desired in the manner described above, and put the time trend for each enterprise back in the total data set. The resulting values will meet the three criteria stated above: The value in a given year for one enterprise keeps its historical relationship to the value in the same year for all other enterprises; all cycles, trends, and natural fluctuations remain intact as occurred in the base period for all enterprises; and the values are in a form which can be changed as desired by the researchers.

After this brief explanation let us turn to a somewhat more detailed explanation of the procedure involved. In general, farm costs (except feed grains and animal replacement costs) are not expected to change cyclically, but in a relatively constant trend over time. Crop yields are not associated with a cyclical behavior, but fluctuate due to weather conditions and have increased in a relatively constant fashion over the last 15-20 years due to technology. A slight cyclic trend may be noticed in some crop prices, but this is probably due largely to the corresponding cyclic fluctuations in livestock prices and the resulting change in demand for feed.

However, in livestock prices, notably hogs and beef cattle, strong cyclic tendencies are readily apparent. Furthermore, over relatively short periods of time (say 20 years) the length and amplitude of these cycles are fairly constant. Since cycles do exist in livestock it is necessary to choose a base period of at least the length of one cycle of the enterprise which has the longest cycle.

Suppose several years of future price and yield data are desired for enterprises such as corn, soybeans, wheat, hogs, sows, steers, and heifers. The approach under this procedure would be as follows:

- (1) Obtain historical data on prices and yields for these enterprises appropriate to the location of interest. Yields may be localized as much as desired (at least to crop reporting districts) and adjusted to reflect desired assumptions regarding cropping practices and abilities of managers.

The price series chosen should reflect the appropriate month or months of possible sale of the product and the market to be used.

(2) Plot the livestock price data to determine the length of the price cycles for the various livestock enterprises.

(3) Choose a base period equal to the length of one or more complete cycles for beef cattle (since beef has the longest cycle). It is also important that one or more whole cycles associated with other livestock enterprises "fit" quite closely within the base period, since when the time trend is removed for each data set it is important that the deflated ends meet. The base period chosen should normally end with the last complete year of available data since the most recent data are usually regarded as the best estimate of future data. The stage of the cycle at the beginning of the base period is largely irrelevant if the cycle is in the same stage at the end of the base period.

(4) Determine the time trend (if any) in prices, yields and other costs associated with each enterprise for the length of the base period. This may be accomplished by simple regression with time or by plotting the data and visually estimating the trend. The time trend should be expressed in units of change per year (i.e. corn yield time trend may be 2.4 bushels increase per year).

(5) Project the data beyond the base period as follows: If there are N years in the base period numbered 1 through N, the value in the first year following the base period is equal to the value in year 1 plus N times the trend per year; the value in the second year following

the base period is equal to the value in year 2 plus N times the trend per year; etc. In more concise notation, let

N = number of years in the base period,

$Y(I,J)$ = yield of enterprise I in year J,

$P(I,J)$ = price of enterprise I in year J, and

$TREND(I)$ = time trend over the base period for enterprise I.

Then $Y^*(I,J+N) = Y(I,J) + N \times TREND(I)$

and $P^*(I,J+N) = P(I,J) + N \times TREND(I)$

Where Y^* and P^* are the new yield and price values sought.

In effect, this step removes the time trend from the data, extends the data, and then replaces the time trend in the whole data set for each enterprise (it is not necessary to physically remove the time trend and replace it, but the effect of doing so is still there.)

Several assumptions are implicit in this approach to extending data for several enterprises. It assumes that relations among enterprises that held true during the base period will be similar in the future. It assumes that weather and other external and internal forces affecting yields will duplicate themselves exactly in the projected data. It assumes that the time trend occurring during the base period will continue beyond the base period. In effect, it assumes that the conditions affecting the chosen enterprises during the base period will continue. Adjustments may have to be made on an individual enterprise basis to reduce the severity of these assumptions. Suppose that for a given enterprise, conditions in the economy become such that its price and/or yield relationship to other enterprises changes or is expected

to change. As soon as this change in structure is foreseen or observed it can be explicitly incorporated into the data.

The approach described is adaptable for the computer to minimize the manual work required. After the data are collected and base period chosen all other steps may be accomplished by the computer.

If none of the enterprises to be considered have price cycles, the described approach would be simplified in that the base period could be of any reasonable length.

ESTABLISHING PRICES AND YIELDS FOR PLANNING PURPOSES

The price and yield environmental simulator just discussed provides data for the levels of product prices and yields which could be used to determine the outcome of a farm over time. It may be useful to obtain another set of data for use by the researcher in planning and decision-making studies over time.

In conditions of imperfect knowledge, decisions are made based on expected or predicted levels of prices and yields, since from the farmer's standpoint, the actual values are unknown. Thus, in certain cases, a set of actual values and a set of predicted values are needed. Let us refer to the set of values to be used in determining the actual outcome of a firm over time as "actual" values and refer to another set of values to be used in planning and decisionmaking for the firm over time as "predicted" values.

Researchers may desire to test the effect of various levels of success in estimating prices and yields since, in a sense, success in forecasting future prices and yields measures the managerial ability of an entrepreneur. Various levels of success in predicting prices

and yields may be compared over time as to their effect on growth of the firm or on some other measure of wellbeing of the firm.

We have developed below a means of establishing values which could be used for determining predicted values or values for planning purposes of a simulated firm which have the following properties: For each actual value a predicted value may be obtained which lies within a given distance (specified by the researcher) from the actual value with a given statistical probability (also specified by the researcher). That is, a formula is derived below which allows the obtaining of a predicted value for each actual value such that the predicted value lies within, for instance, 5% of the actual value, for instance, 90% of the time. The formula will be such that a simple multiplication and addition is conducted in obtaining the predicted values desired.

Equipment needed for this venture includes the following:

- (1) a set of actual values for which corresponding predicted values are desired,
- (2) a decision as to the desired distance from the actual value that the predicted value is to lie (i.e. $\pm 5\%$ of the actual value),
- (3) a decision as to the desired probability of (2) above (i.e. 90%),
- (4) a U^3 value (or t-distribution value with infinite degrees of freedom since $U = t^\infty$) for each probability level desired (i.e. for probability of 90%, $U = 1.6449$),

(5) a population of random normal numbers⁴ with known mean and standard deviation, or access to a computer which can generate the same.

The formula for the general, all-inclusive case, where the distance from the actual value may be any specified distance (D) with any desired probability (U) and where the population of random normal numbers used may have any known mean (\bar{Y}_R) and standard deviation (σ_R), is as follows:

$$Y_e = Y_a [Y_R \cdot D / (U \cdot \sigma_R) + 1 - (\bar{Y}_R \cdot D / (U \cdot \sigma_R))]]$$

where Y_e is the predicted value sought,

Y_a is the actual value known apriori,

Y_R represents an individual observation drawn from the random normal population,

D is the desired deviation from the actual value and must be expressed as a decimal deviation from Y_a (i.e. .05), and

\bar{Y}_R and σ_R are the mean and standard deviation of the random normal population, respectively.

The following formula is for the particular case where Y_e is desired to be within $\pm .05 Y_a$ with a 90% probability and where a population of random normal numbers (such as found in Li) with mean 50 and standard deviation of 10 is used:

$$\begin{aligned} Y_e &= Y_a [Y_R (.05) / (1.6449 \cdot 10) + 1 - (50 (.05) / (1.6449 \cdot 10))] \\ &= Y_a (.00304 Y_R + .848) \end{aligned}$$

Suppose that a predicted value is desired, meeting the above criteria, for an actual value for hogs sold in November 1971 at

\$20.37 per cwt. Suppose further that the number drawn randomly from the table of random normal numbers is 58. Applying the above formula,

$$\begin{aligned} Y_e &= \$20.37 [.00304(58) + .848] \\ &= \$20.37 (1.02432) \\ &= \$20.87 \end{aligned}$$

Before the general formula is derived, additional examples of the application of the general formula to specific cases may be appropriate. Suppose it is desired to obtain predicted values which are within 25% of actual values with 80% probability, and suppose the population of random normal numbers with mean 50 and standard deviation of 10 is used. In this case, $U = 1.2825$. Applying the general formula,

$$\begin{aligned} Y_e &= Y_a [Y_r(.25)/(1.2825 \cdot 10) + 1 - (50(.25)/(1.2825 \cdot 10))] \\ &= Y_a (.01949 Y_r + .0253) \end{aligned}$$

Suppose it is desired to obtain predicted values which are within 25% of the actual values with 80% probability, as above, but with use of a population of random normal numbers of mean zero and standard deviation of one. Applying the general formula,

$$\begin{aligned} Y_e &= Y_a [Y_r(.25)/(1.2825 \cdot 1) + 1 - (0(.25)/(1.2825 \cdot 1))] \\ &= Y_a [.1949 Y_r + 1]. \end{aligned}$$

The derivation of the above general formula is as follows:

The equation for a confidence interval about a value Y_a is $Y_e = Y_a \pm U\sigma$ where U is defined above and σ is the standard deviation of Y_e . The criteria desired is that Y_e should lie the distance $D \cdot Y_a$ from Y_a except for a probability associated with U . Therefore, the

interval $U\sigma$ must equal $D \cdot Y_a$ to fulfill this criteria; or $\sigma = D \cdot Y_a / U$.

Now, if a population of random normal numbers with mean equal to Y_a and standard deviation equal to $D \cdot Y_a / U$ could be found, each observation from it would give a Y_e value meeting the above criteria. A population of this mean and standard deviation can be obtained from a population of random normal numbers with known mean (\bar{Y}_r) and standard deviation (σ_r) by applying two rules⁵ : (1) If a fixed amount is added to or subtracted from each of the observations in a population, the mean will be increased or decreased by that amount, but the variance and standard deviation are not affected. (2) If each of the observations in a population is multiplied by a fixed quantity M , the new mean is M times the old mean and the new standard deviation is M times the old standard deviation.

If each observation in the population is multiplied by the quantity $D \cdot Y_a / (U \cdot \sigma_r)$ the standard deviation would be changed to the desired value $D \cdot Y_a / U$, but the mean would be changed also to the value

$\bar{Y}_r \cdot D \cdot Y_a / (U \cdot \sigma_r)$. To obtain the desired mean of Y_a , the quantity $[1 - (\bar{Y}_r \cdot D / (U \cdot \sigma_r))] \cdot Y_a$ must be added to each observation. This step does not alter the standard deviation. In mathematical terms, the operation performed above is described in the following equation:

$$Y_e = Y_r [D \cdot Y_a / (U \cdot \sigma_r)] + [1 - (\bar{Y}_r \cdot D / (U \cdot \sigma_r))] Y_a$$

or by factoring out Y_a in both right side terms

$$Y_e = Y_a [\bar{Y}_r \cdot D / (U \cdot \sigma_r) + 1 - (\bar{Y}_r \cdot D / (U \cdot \sigma_r))]$$

which is the general formula sought. Once this formula is known, it is not necessary to physically alter the entire population of random

normal numbers by performing the multiplication and addition steps.

The formula is to be applied only to individual observations drawn from this population in obtaining each desired predicted value from the known actual values.

As shown above, the general formula quickly becomes a simple equation for calculating purposes after a population of random normal numbers has been selected, a deviation from the actual values chosen, and a probability attached.

A further statement as to the use of this concept is appropriate. A simple formula such as the one given here may find use in any engineering study in which a set of values is desired to approximate or deviate a given distance from another set of values with a controlled statistical precision. This concept has great potential use in studies involving prices and yields when it is desired to establish predictions based on the actual values but deviating from them in a known and controlled fashion.

In the use of the formula described above, there are no restrictions on the level or the variability of the actual values for which predicted values are desired. Their range may be $-\infty \leq Y_a \leq \infty$.

It may be useful to simulate pessimism or optimism in the obtaining of predicted values by forcing the Y_e values to be less than (in the pessimistic case) or more than (in the optimistic case) the Y_a values. This may be done simply by discarding numbers drawn from the random normal population which are less (greater) than the mean of the random normal population.

Again, this procedure for obtaining a set of predicted values may be computerized to reduce manpower involved if any extremely long list of predicted values is desired.

References

- [1] Harrison, Virden L., Management Strategies and Decision Processes for the Growth of Farm Firms, unpublished Ph.D thesis, Purdue University, August 1970.
- [2] Li, Jerome C.R., Statistical Inference I, Edwards Bros., Inc. Ann Arbor, Michigan, 1964.

Footnotes

* Harrison is Agricultural Economist, FPED, ERS, USDA stationed at Washington, D.C. Morris is Professor of Agricultural Economics, Purdue University.

1. The concepts described here were developed for and used in a farm firm growth study as reported in [1]. The study was a joint effort by USDA and Purdue University.

2. By time trend is meant the change in the data regressed with time. Removing this trend will not alter the cyclic pattern or any other fluctuations in the data.

3. The statistic U follows the normal distribution with mean equal to zero and variance equal to one. For a table of U values see [2, p. 599].

4. For a population of 5000 random normal numbers with mean 50 and standard deviation of 10, see [2, Appendix Table 1].

5. See [2, pp. 9-10], theorems 2.4A and 2.4B.

NATIONAL AGRICULTURAL LIBRARY



1022568168

* NATIONAL AGRICULTURAL LIBRARY



1022568168